UNITED STATES PATENT APPLICATION

OF

PRATHIMA AGRAWAL

DAVID FAMOLARI

TAO ZHANG

FOR

SYSTEM AND METHOD FOR DYNAMICALLY ALLOCATING IP
ADDRESSES FOR SHARED WIRELESS AND WIRELINE NETWORKS
BASED ON PRIORITIES AND GUARD BANDS

SYSTEM AND METHOD FOR DYNAMICALLY ALLOCATING IP ADDRESSES FOR SHARED WIRELESS AND WIRELINE NETWORKS BASED ON PRIORITIES AND GUARD BANDS

5

BACKGROUND OF THE INVENTION

1. FIELD OF THE INVENTION

The present invention generally relates to the use of Mobile Internet protocol (IP) systems, methods and protocols in personal communication systems (PCS) and, more particularly, to a method for dynamically allocating IP addresses for shared wireless and wireline networks based on priorities and guard bands.

2. DESCRIPTION OF THE RELATED ART

With the proliferation of wireless cellular devices (mobile stations), an increasing need has arisen for a methodology and system to provide seamless mobile access to the Internet via existing personal communication services (PCS) networks. To meet this need, a wireless access system must support high-speed packet data transmissions and must also efficiently commuincate with the Internet backbone. While emerging PCS technologies, such as CDMA2000 and Wideband CDMA (WCDMA), define packet data transmission over the air interface, none of these systems present a flexible architecture for connecting a wireless access system to the Internet backbone which is also suitable for general mobile IP interconnectivity.

FIG. 1 is an illustration of a conventional cellular system network 10. The network 10 includes a mobile switching center (MSC) 12 connected to one or more base station controllers (BSC) 14. Each BSC 14 controls one or more base stations (BS) 16 which provide wireless communication links to various mobile stations (MS) 18. The MSC 12 interfaces to a public switched telephone network (PSTN) 20 for routing the voice calls and to a separate signaling system 22 for signaling used for PCS mobility management. Such a conventional cellular system provides Internet access through voiceband modems in the mobile stations, which connect to an Internet access server (not shown) interfaced to the MSC 12, either directly or via the PSTN 20. Alternatively, an interworking function (IWF) 24 interfaced directly to the MSC provides the interworking of protocols between the air interface and the Internet.

Although conventional cellular systems provide adequate data and voice service in some circumstances, such network configurations suffer from several deficiencies. In particular, the connection model is based on dialup modems where the IWF acts as a server, providing a point-to-point connection to the MS and assigning a temporary IP address to the

5   MS that is valid only during the session. When the MS changes the point of attachment with the Internet (i.e., a hand off is performed), a new session must be established and a new temporary address must be assigned to the MS using an address assignment scheme. Therefore, an Internet session cannot be maintained as the MS changes the point of attachment.

10   As the MS moves across different subnets, packet delivery to the MS is based on the temporary IP address and not the static IP address of the MS. Since the MS does not have a static IP address, datagrams cannot be delivered to the MS until the MS has established a Point-to-Point Protocol (PPP) session and the host attempting to send data knows the temporary IP address of the MS.

15   The advent of wireless IP networks has presented challenging problems in the area of dynamic IP address allocation. Wireless IP networks comprise a plurality of network cells wherein a plurality of handoff and resident hosts communicate with the wireless IP network. Handoff hosts are network hosts whose initial network connection was established in another cell and that subsequently migrated to a new network cell. In contrast, resident

20   hosts are hosts whose initial network connection was established in the present cell and that have not migrated to another network cell. Each network cell further includes a base station that maintains handoff and resident host network connections within the network cell, and facilitates dynamic handoff of handoff hosts within its cell.

The base station within each cell facilitates a network connection with its

25   handoff and resident users through IP addresses assigned to the network hosts within the cell. Thus, an IP address must be allocated to the handoff and resident hosts within a cell to establish and maintain a network connection. The IP addresses dedicated to the handoff and resident hosts can be allocated by methods including an IP address server system(s) that assigns IP addresses for new hosts, and base stations themselves that may determine their own

30   necessary IP address allocation. IP address servers that allocate IP addresses can include servers that are part of a base station itself, or can include IP address servers that are attached

to the network and send IP address allocations to the base station within a cell. These IP address allocation methods and apparatuses confront a number of problems when allocating IP addresses to handoff and resident hosts within their cells.

First, whenever a host initially establishes a network connection within a cell, the host is considered a resident host, because it establishes its initial network connection within the cell. In order for the resident host to establish its network connection, a base station within the cell must allocate an IP address from the cell's IP address pool to the resident host. In contrast, whenever a resident host migrates to a new cell after already having its network connection in a prior cell, the host becomes a handoff host, and the new cell must maintain the preexisting network connection of the handoff host. If the new cell has a different IP subnet address relative to the prior cell, the new cell must allocate a new IP address to the handoff host in order to maintain the handoff host's network connection from the prior cell.

Current dynamic IP address allocation protocols such as Dynamic Host Configuration Protocol (DHCP) treat resident and handoff hosts as equivalent and without preference when allocating IP addresses to these hosts. Thus, a base station may allocate an IP address to a resident host instead of a handoff host, thereby permitting a new resident network connection in place of the prior network connection of a handoff host network, which then must be dropped. This equal treatment is obviously undesirable, because denying a resident host an IP address will merely prevent a new network connection, whereas denying a handoff host an IP address causes its preexisting network connection to be terminated. Thus, IP address allocations to handoff hosts should be preferred to IP address allocation of resident hosts in order to facilitate the handoff of preexisting handoff host network connections.

Second, the problem of handoff and resident host network access is further aggravated due to the fact that there is no minimum number of IP addresses that are reserved by individual cells to ensure a certain level of host access. The minimum number of IP addresses available will vary from cell to cell, and thus handoff and resident hosts will have differing degrees of access to the IP network based on the particular cell wherein the host is communicating with the network. As a result, fair access to the wireless IP network for handoff and resident hosts differs from cell to cell, because one cell may have significantly fewer IP addresses to allocate relative to another cell. This is another limitation of present IP

address allocation protocols, because it is desirable to ensure fair access of handoff and resident hosts to the network regardless of the particular cell wherein these hosts reside.

In sum, present dynamic IP address allocation protocols for wireless IP network systems suffer from two significant problems. First, these protocols treat resident and handoff hosts equally when allocating IP addresses, thereby causing preexisting handoff host network connections to be dropped in favor of new resident host network connections, when it is preferable to preserve a preexisting handoff host network connection by denying a new resident host network connection. Second, there is no minimum number of IP addresses allocated to individual cells, thereby causing fair access problems when different cells provide different numbers of IP addresses for handoff and resident hosts. As a result, network access is arbitrarily dependent upon the cell in which the handoff and network host reside.

Accordingly, it is apparent that there is a need and desire for a method for dynamically allocating IP addresses to wireless and wireline IP hosts that share the same pool of IP addresses such that handoff hosts will possess a higher probability of obtaining a new IP address than recently powered local hosts, and such that a proper balance is achieved between the number of IP addresses allocated to wireless and wireline IP hosts.

SUMMARY OF THE INVENTION

The present invention is a system and method for dynamically allocating IP addresses for shared wireless and wireline networks based on priorities and guard bands. In accordance with the invention, the method differentiates between wireless hosts and wireline hosts such that priorities for connecting each host are established prior to handing off mobile calls.

A single system IP address server is used to periodically and dynamically update allocation pools of IP addresses which are reserved for wireless or wireline hosts. The system IP address server controls an entire address pool, contains predictive analysis logic, performs a predictive analysis, partitions the IP addresses into distinct address pools, and maintains "wireless guard bands" of IP addresses for wireless users to ensure that a higher priority address pool is never exhausted before an address pool which has a lower priority. The system IP address server also possesses an interface with which to add or delete IP addresses from each of the distinct address pools as required.

Within the pool of IP addresses allocated to wireless hosts, a "handoff guard band" of IP addresses is established for handoff hosts. Further, handoff hosts are provided with a higher access priority than other hosts such that system IP address servers process address requests from the handoff hosts before processing requests from other hosts. As a

5   result, handoff latencies for wireless calls are reduced.

Predictive analysis is performed in the system IP address server to determine the number of IP addresses to allocate to wireless terminals in a handoff state, wireless terminals in a resident state, and wired terminals. In certain embodiments, the predictive analysis is performed using a moving weighted average, where the average number of requests

10   from hosts in each category is recorded and the average number of total IP addresses over a suitable fixed period of time is computed. In other embodiments, the size of the guard bands is calculated statistically by determining the required volume of IP addresses over a fixed period of time. Here, various statistical parameters may be used, such as the time of day, the "mix" of traffic, special situations (expected and unexpected), or the like.

15   A priority is assigned to a host requesting a new IP address. Handoff hosts are provided with a higher access priority than other hosts such that system IP address servers process address requests from the handoff hosts before processing requests from other hosts. Priorities of the hosts are communicated to the IP address servers (e.g., the DHCP servers). As a result, the servers will first service the host having the highest priority. This permits

20   wireless terminals to obtain IP addresses faster during handoff, and therefore reduces handoff latencies.

A classification process is used to classify wireless hosts into handoff hosts and dormant hosts. Handoff hosts carry active sessions across network boundaries and are rapidly configured with a valid IP address to preserve application level quality. Dormant hosts are

25   those that are not actively communicating, and therefore possess greater latency latitudes with respect to the configuration of the IP address.

A minimum number of addresses, i.e., a wireless handoff IP address pool or a wired terminal IP address pool (a classification), which cannot be affected by a particular reallocation scheme, is made available to each category of device accessing the network. As a

30   result, the system IP address server will be unable to revoke addresses from a classification such that the total number of addresses available in the classification is never brought below

the guard band number. In addition, if a number of addresses are allocated as a result of wireless terminal requests that brings the total available IP addresses in such an allocation pool below the number of addresses in the guard band, an automatic request occurs to replenish the allocation pool up to the guard band level.

5   In accordance with the invention, a wireless handoff host is distinguished from a wireless resident host by determining whether the host has changed, or is in the process of changing, its current point of attachment to the network. For instance, handoff hosts are required to have undergone link-layer handoff procedures to establish a valid link layer with which to communicate with a base station. The base station will possess knowledge of this,

10   and can use this information to identify hosts that have been handed off from other base stations.

A wireless IP address agent resides in the wireless network. Each wireless IP address agent handles requests for IP addresses from the wireless terminals, categorizes the wireless terminal as handoff or resident, and then forwards the IP address associated with the

15   handoff or resident to the system IP address server.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and other advantages and features of the invention will become more apparent from the detailed description of the preferred embodiments of the invention

20   given below with reference to the accompanying drawings in which:

FIG. 1 is an illustration of a conventional cellular system network;

FIG. 2 is a schematic block diagram of shared IP servers without base station requests;

FIG. 3 is a schematic block diagram of shared IP servers with base station

25   requests;

FIG. 4 is a schematic block diagram illustrating base stations with a dedicated IP address server in accordance with the invention; and

FIG. 5 is a flow chart illustrating the steps of an embodiment of the method of the invention for dynamically allocating Internet protocol addresses for shared wireless and

30   wireline networks.

DETAILED DESCRIPTION OF EXEMPLARY EMBODIMENTS

FIG. 2 is a schematic block diagram of shared IP servers without base station requests in accordance with an exemplary embodiment of the present invention. The system and method of the invention provides dynamic addressing of servers (e.g., DHCP servers)

5   such that the allocation of a proper portion of IP addresses to wireless host and wireline hosts is ensured. A wireless host **220** is an end user device which is connected to a network **290** via wireless channels. Such devices can be further categorized by whether they are in a handoff state (handoff) or are residents of a current wireless base station (resident).

In accordance with the invention, a single system IP address server **240** is used

10 to periodically and dynamically update allocation pools of IP addresses which are reserved for wireless or wireline hosts **220**. The system IP address server **240** controls an entire address pool **250**, contains predictive analysis logic **260**, performs a predictive analysis, partitions the IP addresses **250** into distinct address pools **270, 275, 280**, such that "wireless guard bands" of IP addresses for wireless users are maintained to ensure that a higher priority address pool is

15 never exhausted before an address pool which has a lower priority. The system IP address server **240** also possesses an interface (not shown) with which to add, or delete, IP addresses from each of the distinct address pools **270, 275, 280** as required. A guard band is a pool of IP addresses which are dedicated to a specific type of host (e.g., wireless hosts). Within the pool of IP addresses allocated to wireless hosts, a "handoff guard band" of IP addresses **270** is

20 established for handoff hosts, a "resident guard band" of IP addresses **275** is established for resident hosts, and a "wired guard band" of IP addresses **280** is established for wired terminals.

The predictive analysis **260** is performed in the system IP address server **240** to determine the number of IP address to allocate to wireless terminals in a handoff state,

25 wireless terminals in a resident state, and wired terminals. In certain embodiments, the predictive analysis is performed using a moving weighted average, where the average number of requests from hosts in each category is recorded and the average number of total IP addresses over a suitable fixed period of time is computed. In the preferred embodiment, the suitable fixed period of time is approximately 10 minutes. In other embodiments, the size of

30 the guard bands is calculated statistically by determining the required volume of IP addresses

over a fixed period of time. Here, various statistical parameters may be used, such as the time of day, the "mix" of traffic, special situations (expected and unexpected), or the like.

The following IP address guard bands are established within the system IP address server **240**:

5   (i) *Wireless guard bands*: A pool of IP addresses is dedicated to wireless hosts. The size of the wireless guard band is dynamically based on the changing IP address demands of wireless and wireline terminals in the domain of the IP address servers (e.g., the DHCP servers); and

       (ii) *Handoff guard bands*: A pool of IP addresses from the IP space is allocated 10 to wireless users for the support of handoff hosts. The size of such "handoff guard bands" is dynamically adjusted based on the handoff volume. IP address servers **240** dynamically predict the size of each guard band. The IP address servers then partition the shared pool of IP addresses accordingly before the anticipated IP address demands actually arrive. The shared address server **240** predicts the sizes of the address guard bands for each base station, and 15 logically partitions the available shared address space on the system IP address server **240** for the guard bands based on information gathered from each base station.

A priority is assigned to a host requesting a new IP address. Handoff hosts are provided with a higher access priority than other hosts, such that system IP address servers process address requests from the handoff hosts before processing requests from other hosts. 20 Priorities of the hosts are communicated to the IP address servers (e.g., the DHCP servers). As a result, the servers will service the host having the highest priority first. This permits wireless terminals to obtain IP addresses faster during handoff, and therefore reduces handoff latencies.

A classification process **210a**, **210b** is used to classify wireless hosts into 25 handoff hosts and dormant hosts. Handoff hosts carry active sessions across network boundaries and are rapidly configured with a valid IP address to preserve application level quality. Dormant hosts are those that are not actively communicating, and therefore possess greater latency latitudes with respect to configuration of the IP address.

A minimum number of addresses or guard band in the wireless handoff IP 30 address pool **270**, the resident handoff IP address pool **275** and the wired terminal IP address pool **280** (a classification) are set aside so that each are unaffected by a particular reallocation

scheme. As a result, the system IP address server **240** will be unable to revoke addresses from a classification such that the total number of addresses available in the classification is never brought below the guard band number. In addition, if a number of addresses are allocated as a result of wireless terminal requests that brings the total available IP addresses in such an allocation pool below the number of addresses in the guard band, an automatic request occurs to replenish the allocation pool up to the guard band level. For example, if the number of IP addresses assigned to a "wireless resident" host address pool is 100 with a guard band of 10, then in accordance with the invention at least 10 addresses are maintained in the pool. Hence, if 95 "wireless resident" hosts addresses are requested and granted, an automatic attempt is made to allocate at least another 5 addresses to this category to maintain the appropriate guard band.

In accordance with the invention, a wireless handoff host is distinguished from a wireless resident host by determining whether the host has changed, or is in the processes of changing, its current point of attachment to the network. For instance, handoff hosts are required to have undergone link-layer handoff procedures to establish a valid link layer with which to communicate with a base station (not shown). The base station will possess knowledge of this, and can use this information to identify hosts that have been handed off from other base stations.

Implementation of the classification process **210a**, **210b** varies according to the specific wireless access technology in use, such as CDMA, Time Division Multiple Access (TDMA), 802.11b, or the like. When using CDMA, a handoff is always preceded by a *Handoff Request Message*. This provides the base station with the ability to determine whether a terminal is a wireless handoff terminal based on the received *Handoff Request Message*. This is not the case with wireless resident hosts, and there are no handoff messages received by the base station associated with these terminals. Therefore the classification process can use this information to determine the nature of the request. If a request comes from a wireless terminal "shortly" after a *Handoff Request Message* was received from this terminal, the wireless terminal is classified as a wireless handoff host, otherwise it is classified as a wireless resident host. Here, "shortly" is implementation specific, and is programmed by the network administrator such that it is any time period desired. However, a time period which is too large is disadvantageous when used in the present contemplated embodiments. In

-9-

the preferred embodiment, a time period in the range of approximately 30 to 60 seconds is used. Hence, when an address request is received from a terminal having a prior *Handoff Request Message* associated with it in the previous 60 seconds, the terminal is classified as a wireless handoff host.

5   Wireless IP address agents **200a**, **200b** are software which resides in the wireless network **290**. Each wireless IP address agent handles requests for IP addresses from the wireless terminals **220**, categorizes the wireless terminal as handoff or resident, and then forwards the IP address associated with the handoff hosts or resident hosts to the system IP address server **240**. In embodiments, the wireless IP address agents **200a**, **200b** reside in

10 wireless base stations (not shown) which provide link layer attachments to wireless terminals. Here, the base stations receive IP address requests and forward them to the system IP address server **240**. As a result, the system IP address server **240** can distinguish between IP address requests which originate from wireless or wireline terminals.

         FIG. 3 is a schematic block diagram of shared IP servers with base station

15 requests from mobile terminals **220**. Here, each base station **330**, **350**, **370** uses a predictive analysis **340**, **360**, **380** to predict a required guard band size for an IP address pool **335**, **355**, **375**. The results of the prediction are provided to a shared system IP address server **300**. A final decision on the guard band sizes for each base station is then performed in the system IP address server **300** based on the predictions from the base stations and other overall system

20 considerations, such as the minimum number of addresses to achieve a fair chance of access from each cell.

         FIG. 4 is a schematic block diagram illustrating base stations having a dedicated IP address server in accordance with the invention which respond to requests for IP addresses from mobile terminals **402**. In this configuration, a system IP address server is not

25 used, i.e., there is no central authority. Instead, all base stations **400**, **430**, **450** posses an IP address server **415**, **440**, **460** wherein a respective predictive analysis **420**, **445**, **465** is performed to determine a required guard band size for a respective IP address pool **410**, **435**, **455**. The base stations coordinate address allocations during base station-to-base station intercommunication.

30   FIG. 5 is a flow chart illustrating the steps of an embodiment of the method of the invention for dynamically allocating Internet protocol addresses for shared wireless and

wireline networks. In accordance with the invention, a determination of a total IP address pool requirement is performed using a predictive analysis, as indicated in step **500**. In embodiments of the invention the initial allocation is configured by a network administrator. In addition, the predictive analysis may be performed using a number of predictive algorithms, such as a moving weighted average. Here, the average number of requests from hosts in each device category is recorded, and the number of requests over a fixed period of time is calculated. In the preferred embodiment, an exemplary time period is 10 minutes.

Next, the total IP address pool is partitioned into groups of addresses which are designated for use with an associated user group or classification, i.e. a handoff group, a resident group and/or a wired terminal group, as indicated in step **510**.

An IP address server is used to monitor IP address demands associated with a cell, as indicated in step **520**. A check is performed to determine whether an update of the total IP address pool is necessary, as indicated in step **530**. If this is the case, the IP address server will periodically and dynamically update the allocations reserved for the groups of addresses which are designated for use with the associated user group, as indicated in step **540**. A return to step **520** occurs to continue monitoring IP address demands associated with the cell. If no update is necessary, the method of the present invention will return to step **520** where monitoring of IP address demands associated with the cell is repeated. In this manner, continuous monitoring of IP address demand is performed to ensure the continuous allocation of a proper portion of IP addresses to wireless host and wireline hosts.

A number of features of the present invention will be apparent to one skilled in the art. First, third parties such as wireless Internet Service Providers (ISPs) can control allocation of IP addresses to network cells. The present invention enables wireless and wireline ISPs to specify and control how IP addresses should be allocated to network cells in the wireless IP network, thereby allowing the ISP to allocate more IP addresses in specific network cells. For instance, an ISP may allocate more IP addresses to an area where high-paying customers reside so that network performance in that area is enhanced.

When allocating IP addresses to cells, third-party ISPs may directly specify the number of IP addresses to allocate to each cell or geographical area that is covered by multiple network cells in order to support the third-party's services. A third-party may also specify performance objectives that can be achieved by a plurality of IP address allocations to the

network cells, and then allow the IP address servers to allocate IP addresses to meet these performance objectives. For instance, a third-party may specify call blocking probabilities and quality of service for a geographical area, and then allow the wireless IP network address servers to determine the actual IP address allocations for cells within the geographical area to

5   meet these performance objectives.

Using the method of the invention, a rapid priority based configuration of IP addresses for wireless users is achieved. IP addresses are reserved for handoff hosts such that handoff call blocking due to lack of IP addresses in a new cell is minimized. In addition, the reservation of IP addresses for wireless dormant hosts minimizes the time to configure IP

10   addresses.

Although the invention has been described and illustrated in detail, it is to be clearly understood that the same is by way of illustration and example, and is not to be taken by way of limitation. The spirit and scope of the present invention are to be limited only by the terms of the appended claims.

15